

GenCore version 5.1.3
Copyright (c) 1993 - 2002 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: December 13, 2002, 20:59:48 ; Search time 2883 Seconds

(without alignments)
1655.519 Million cell updates/sec

Title: US-09-659-737A-1

Perfect score: 164

Sequence: 1 ggcacgggacatcaagagcag.....tgccctgatgccccagaaag 164

Scoring table: IDENTITY_MDC

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

Listing first 45 summaries

GenBml:*

1: gb_ba:*

2: gb_htg:*

3: gb_in:*

4: gb_ov:*

5: gb_ov:*

6: gb_pat:*

7: gb_ph:*

8: gb_pl:*

9: gb_pr:*

10: gb_ro:*

11: gb_sy:*

12: gb_un:*

13: gb_un:*

14: gb_vi:*

15: em_ba:*

16: em_fun:*

17: em_hum:*

18: em_in:*

19: em_mu:*

20: em_or:*

21: em_ov:*

22: em_ov:*

23: em_pat:*

24: em_ph:*

25: em_pl:*

26: em_ro:*

27: em_sts:*

28: em_un:*

29: em_vi:*

30: em_htg_hum:*

31: em_htg_inv:*

32: em_htg_other:*

33: em_htg_mus:*

34: em_htg_plg:*

35: em_htg_rtd:*

36: em_htg_mam:*

37: em_htg_vrt:*

38: em_sy:*

39: em_htgo_hum:*

40: em_htgo_mus:*

41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	164	100.0	164	6	AX139125 Sequence
2	152.8	93.2	2157	6	AX207410 Sequence
3	152.8	93.2	3111	6	AX482007 Sequence
4	152.8	93.2	3518	6	AX482009 Sequence
5	149.6	91.2	3910	9	AX311797 Homo sapi
6	149.6	91.2	4667	9	AX311798 Homo sapi
7	136.4	83.2	118632	9	AX133380 Human DNA
8	120.8	73.7	3616	10	BC021891 Mus muscu
9	106	64.6	168795	2	AC102420 Mus muscu
10	104.4	63.7	96101	2	AC112003 Rattus no
11	94.2	57.4	3931	9	AF251442 Homo sapi
12	83	50.6	3531	9	HSU07747 Human SH3 d
13	83	50.6	3558	6	AX399680 Sequence
14	83	50.6	3558	9	HOMMLK3A Homo prote
15	83	50.6	3603	9	BC011263 Homo sapi
16	83	50.6	3799	9	AK092015 Homo sapi
17	82	50.0	166200	9	AC004816 Homo sapi
18	79.2	48.3	39251	5	AF154413 Fugu rubr
19	78.8	48.0	197943	2	AC125351 Mus muscu
20	78.8	48.0	203574	2	AC124595 Mus muscu
21	75	45.7	2390	10	BC030928 Mus muscu
22	73.4	44.8	3138	9	HSMTWR X90846 H. sapiens m
23	73.4	44.8	3454	6	AX337846 Sequence
24	72.8	44.4	150694	2	AP001459 Homo sapi
25	72.8	44.1	686	11	G33125 Ep10F7a5 Hu
26	72.4	44.1	140356	2	AP000803 Homo sapi
27	72.4	44.1	162445	2	AC022488 Homo sapi
28	72.4	44.1	211382	9	AP001362 Homo sapi
29	72.4	44.1	211382	9	AP001362 Homo sapi
30	72.4	44.1	211382	9	AP001362 Homo sapi
31	67.8	41.3	66251	2	AC100718 Mus muscu
32	66	40.2	34340	2	AC130251 Rattus no
33	64.8	39.3	168423	2	AC111962 Rattus no
34	64.4	38.0	46440	10	AF151142 Mus muscu
35	62.4	37.8	137	6	AF151142 Mus muscu
36	62	36.5	215615	2	AC074312 Mus muscu
37	60	36.5	226059	2	AC074312 Mus muscu
38	60	36.5	243275	2	AC079488 Mus muscu
39	60	36.5	243275	2	AC073705 Mus muscu
40	59.8	36.5	140529	2	AC011486 Homo sapi
41	59.8	36.5	178335	2	AC118344 Homo sapi
C 41	58.4	35.6	170094	2	AC120811 Rattus no
C 42	58.4	35.6	170094	2	AC120811 Rattus no
C 43	57.4	35.0	3447	3	AF416233 Drosophila
44	57.4	35.0	5129	3	AY045717 Drosophila
45	57.4	35.0	5131	3	AY119549 Drosophila

ALIGNMENTS

RESULT 1

AX139125

LOCUS AX139125 164 bp DNA linear PAT 30-MAY-2001

DEFINITION Sequence 1 from Patent EP1085093.

ACCESSION AX139125

VERSION AX139125.1 GI:14274801

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE 1 (bases 1 to 164)

AUTHORS Blumenberg, M. and Gazel, A.M.

TITLE Genes and polynucleotides associated with ultraviolet radiation-mediated skin damage and uses thereof

JOURNAL Patent: EP 1085093-A 1 21-MAR-2001;
NEW YORK UNIVERSITY (US)

FEATURES
source location/Qualifiers
1..164
/organism="Homo sapiens"
/db_xref="taxon:9606"
<2..>163
/note="unnamed protein product"

CDS

/codon_start=1
/protein_id="CAC3965.1"
/db_xref="GI:14274802"
/translation="HRDIKAGNILLLEKIEHDICNKTITPDGLAREHRTTKMST
AGTYAMWAP"

BASE COUNT 55 a 33 c 44 g 32 t

ORIGIN

Query Match 100.0%; Score 164; DB 6; Length 164;
Best Local Similarity 100.0%; Pred. No. 4,4e-38;
Matches 164; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCACCGGGACATCAAGCGAGGAATATTTGCTACTTGAGAGATAGAACATGATGACAT 60
|||||
Db 1 GCACCGGGACATCAAGCGAGGAATATTTGCTACTTGAGAGATAGAACATGATGACAT 60

QY 61 CTGCAATAAACCTTTGAAGATTACAGATTTGGGTGGCGAGGAATGGCACAGACCAC 120
|||||
Db 61 CTGCAATAAACCTTTGAAGATTACAGATTTGGGTGGCGAGGAATGGCACAGACCAC 120

QY 121 CAAATATGAGCAGCAGCAGCACTATGCTGATGGCCCCCAGAG 164
|||||
Db 121 CAAATATGAGCAGCAGCAGCACTATGCTGATGGCCCCCAGAG 164

RESULT 2
AX207410 2157 bp DNA linear PAT 30-AUG-2001

LOCUS AX207410
DEFINITION Sequence 23 from Patent WO0155356.
ACCESSION AX207410
VERSION AX207410.1 GI:15395228
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
TITLE 1 (bases 1 to 2157)
PLOWMAN,G., Whyte,D., Manning,G., Sudarsanam,S. and Martinez,R.
HUMAN protein kinases and protein kinase-like enzymes
JOURNAL Patent: WO 0155356-A 23 02-AUG-2001;
Sugen, Inc. (US)

FEATURES
source location/Qualifiers
1..2157
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 514 a 611 c 638 g 394 t

ORIGIN

Query Match 93.2%; Score 152.8; DB 6; Length 2157;
Best Local Similarity 95.7%; Pred. No. 1.1e-34;
Matches 157; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GCACCGGGACATCAAGCGAGGAATATTTGCTACTTGAGAGATAGAACATGATGACAT 60
|||||

Db 780 GCACCGGGACCTCAAGTCCGCAACATTTGCTACTTGAGAGATAGAACATGATGACAT 839
|||||

QY 61 CTGCAATAAACCTTTGAAGATTACAGATTTGGGTGGCGAGGAATGGCACAGACCAC 120
|||||
Db 840 CTGCAATAAACCTTTGAAGATTACAGATTTGGGTGGCGAGGAATGGCACAGACCAC 899
|||||

QY 121 CAAATATGAGCAGCAGCAGCACTATGCTGATGGCCCCCAGAG 164
|||||
Db 900 CAAATATGAGCAGCAGCAGCACTATGCTGATGGCCCCCAGAG 943
|||||

RESULT 3
AX482007 3111 bp DNA linear PAT 16-AUG-2002

LOCUS AX482007
DEFINITION Sequence 1 from Patent WO02055685.
ACCESSION AX482007
VERSION AX482007.1 GI:22316731
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Hu,Y., Kieke,J.A. and Donoho,G.
TITLE Novel human kinase and polynucleotides encoding the same
JOURNAL Patent: WO 02055685-A 1 18-JUL-2002;
Lexicon Genetics Incorporated (US)

FEATURES
source location/Qualifiers
1..3111
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 740 a 905 c 864 g 594 t 8 others

ORIGIN

Query Match 93.2%; Score 152.8; DB 6; Length 3111;
Best Local Similarity 95.7%; Pred. No. 1.1e-34;
Matches 157; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GCACCGGGACATCAAGCGAGGAATATTTGCTACTTGAGAGATAGAACATGATGACAT 60
|||||
Db 780 GCACCGGGACCTCAAGTCCGCAACATTTGCTACTTGAGAGATAGAACATGATGACAT 839
|||||

QY 61 CTGCAATAAACCTTTGAAGATTACAGATTTGGGTGGCGAGGAATGGCACAGACCAC 120
|||||
Db 840 CTGCAATAAACCTTTGAAGATTACAGATTTGGGTGGCGAGGAATGGCACAGACCAC 899
|||||

QY 121 CAAATATGAGCAGCAGCAGCACTATGCTGATGGCCCCCAGAG 164
|||||
Db 900 CAAATATGAGCAGCAGCAGCACTATGCTGATGGCCCCCAGAG 943
|||||

RESULT 4
AX482009 3518 bp DNA linear PAT 16-AUG-2002

LOCUS AX482009
DEFINITION Sequence 3 from Patent WO02055685.
ACCESSION AX482009
VERSION AX482009.1 GI:22316732
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens

REFERENCE
AUTHORS Hu,Y., Kieke,J.A. and Donoho,G.
TITLE Novel human kinase and polynucleotides encoding the same
JOURNAL Patent: WO 02055685-A 3 18-JUL-2002;
Lexicon Genetics Incorporated (US)

FEATURES
source location/Qualifiers
1..3518
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 811 a 1043 c 976 g 680 t 8 others

ORIGIN

Query Match 93.2%; Score 152.8; DB 6; Length 3518;
Best Local Similarity 95.7%; Pred. No. 1.1e-34;
Matches 157; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 1 GCACCGGGACATCAAGCGAGGAATATTTGCTACTTGAGAGATAGAACATGATGACAT 60
|||||
Db 1043 GCACCGGGACCTCAAGTCCGCAACATTTGCTACTTGAGAGATAGAACATGATGACAT 1102
|||||

QY 61 CTGCAATAAACCTTTGAAGATTACAGATTTGGGTGGCGAGGAATGGCACAGACCAC 120
|||||

Db	1103	CTGCATTAACCTTTGAAGATTACAGATTCTTGGGTGGCGAGGCAATGGCACAGGACCAC	1162
QY	121	CAAAATGAGCAGCAGGACACCTATGCTTGATGGCCCCAGAG	164
Db	1163	CAAAATGAGCAGCAGGACACCTATGCTTGATGGCCCCAGAG	1206
RESULT 5			
LOCUS	HSA311797	3910 bp	linear
DEFINITION	HSA311797		PRI 12-DEC-2001
ACCESSION	AJ311797		
VERSION	AJ311797.1	GI:17736728	
KEYWORDS	mixed lineage kinase 4alpha; MLK4ALPHA gene.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.		
AUTHORS	Kvashe, S., Protolopov, A., Rynditch, A., Zabarovsky, E. and Kashuba, V.		
TITLE	MLK4, a new member of mixed lineage kinases		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 3910)		
AUTHORS	Kashuba, V.		
TITLE	Direct Submission		
JOURNAL	Submitted (12-APR-2001) Kashuba V., Microbiology and Tumorbiology Center, Karolinska Institute, Box 280, 171 77, SWEDEN		
FEATURES	Location/Qualifiers		
source	1..3910		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/chromosome="1"		
	/map="1q42"		
	/tissue_type="heart"		
	262..1974		
	/gene="MLK4ALPHA"		
	262..1974		
	/gene="MLK4ALPHA"		
	/codon_start=1		
	/evidence="experimental"		
	/product="mixed lineage kinase 4alpha"		
	/protein_id="CAC84638.1"		
	/db_xref="GI:17736729"		
	/translation="MALRGAAGAADTPVSSAGAPGGSSSSSTSSGGSASAGAGLWNAALYFEARGEDELRLRQLEVLSSODAAVSGDGSWMAQVORLGIFFPANYVAPCRERAAAPRPSPSPVHVAFERLELKEKELGAGFGQYVATWQGOEVAVKARQPDREAAASAEVRRARRLEPAMLRHNITIELNGCYLOOHCIVLEFARGALNRLAANAANAADPAPRPRRRARRIPHYLVWMAVQIARGMLTYHEBAPVPLIHLDKSSNILLIEEIEHDDICNKTLLITDFGLANRHRRTKMSYAGVAMAPAVIKSLPSKSDIWSGCVLLWELLTGEVPRIDGLAVAVYAVANKLTLPLPSTCPPEFAPALMECWODPHIRPSFRALELILQUTLIEGAVNTEMPOESFHSMODMKLEIQMEDELTKEREKLSREELTAAALLOQSOEELKLRQDLAEREIDVLERLNTLITOLNDEKPKVKKRGRKFSRLKIKLKDGHILSLPSDFQHKITTYQASPNLDKRSLSNSSSSPSSPTMRLRLAIQCELSALPPRGLLC"		
BASE COUNT	polyA signal 923 a 1038 c 1056 g 893 t		
ORIGIN			
Query Match	91.2%;	Score 149.6;	DB 9; Length 3910;
Best Local Similarity	94.5%;	Pred. No. 1e-33;	
Matches 155;	Conservative	0; Mismatches	9; Indels
		Gaps	0;
QY	1	GCACGGGACATCAAGGCGAATAATTTTCTACTTGAGAGATGAACATGATGACAT	60
Db	1041	GCACGGGACCTCMACTCCAGCACATTTTCTACTTGAGAGATGAACATGATGACAT	1100
QY	61	CTGCATTAACCTTTGAAGATTACAGATTCTTGGGTGGCGAGGGAATGGCACAGACCA	120
Db	1101	CTGCATTAACCTTTGAAGATTACAGATTCTTGGGTGGCGAGGGAATGGCACAGACCA	1160
QY	121	CAAAATGAGCAGCAGGACACCTATGCTTGATGGCCCCAGAG	164
Db	1161	CAAAATGAGCAGCAGGACACCTATGCTTGATGGCCCCAGAG	1204

[illegible]

|||||

Db 1161 CAAATGACACAGCAGCATTGCTGATGGCCCCGAG 1204

RESULT 7
AL133380
LOCUS
DEFINITION

AL133380 118632 bp DNA linear PRI 29-NOV-2000
Human DNA sequence from clone RP5-862P8 on chromosome 1q42.2-43
contains a 40S ribosomal protein S7 pseudogene, start of a gene for
a novel protein (similar to C.elegans and D.melanogaster), a gene
similar to MAP3K10 (mitogen-activated protein kinase kinase
10), STSs, GSSs and a Cpg island, complete sequence.
AL133380
AL133380.5 GI:8217443
HTG: Cpg island; protein kinase; Ribosomal protein.
human.
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 118632)
Hall, R.
Direct Submission
Submitted (31-MAY-2000) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jun 3, 2000 this sequence version replaced gi:7798842.
During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence has been finished according to sequence map criteria
as follows. An attempt is made to resolve all sequencing problems,
such as compressions and repeats, but not necessarily within known
annotated human repeat sequence elements (e.g. Alu). Where the
sequence is ambiguous, there is an annotation using the 'unsure'
feature key.
The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence
was generated from part of bacterial clone contigs of human
chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping
Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chrl
RP5-862P8 is from the library RPCI-5 constructed at the Roswell
Park Cancer Institute by the group of Pieter de Jong. For further
details see http://bacpac.med.buffalo.edu/
VECTOR: pCYPAC2
This sequence is the entire insert of clone RP5-862P8.
FEATURES
source
1.118632
Location/Qualifiers
1..118632
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="1"
/map="q42.2-43"
/clone="RP5-862P8"
/clone_lib="RPCI-5"
323..682
repeat_region
/note="RFLB repeat: matches 1..362 of consensus"
742..953
repeat_region
/note="MIR repeat: matches 45..262 of consensus"
misc_feature
complement(1784..2072)
/note="match: GSS: Em:AQ887379"
misc_feature
complement(1793..1994)
/note="match: GSS: Em:AQ384546"
1793..1850
repeat_region
/note="29 copies 2 mer tt 74% conserved"
misc_feature
complement(1810..2413)
/note="match: STS: Em:D11774"
misc_feature
complement(1812..1928)

gene
CDS
misc_feature
1873..2009
/evidence=not_experimental
/note="match: GSS: Em:AQ219188"
complement(1849..2433)
/gene="dJ862P8.1"
misc_feature
1875..2308
/note="match: GSS: Em:AQ476700"
1875..2308
/note="match: GSS: Em:AQ593042"
2153..2376
/note="match: GSS: Em:AQ264373"
2157..2439
/note="match: GSS: Em:AQ554409"
2186..2451
/note="match: STS: Em:016258"
complement(2283..2360)
/gene="dJ862P8.1"
misc_feature
2841..2910
/note="match: STS: Em:H84775"
/note="MER91A repeat: matches 1..69 of consensus"
4083..4220
/note="MIR repeat: matches 2..142 of consensus"
4207..4313
/note="MIR repeat: matches 62..173 of consensus"
4314..4623
/note="Alusp repeat: matches 1..312 of consensus"
5181..5494
/note="Alusx repeat: matches 1..312 of consensus"
complement(6410..6697)
/note="match: GSS: Em:AQ088034"
complement(6422..6787)
/note="match: GSS: Em:AQ177047"
6688..6884
/note="MIR repeat: matches 1..247 of consensus"
6974..7178
/note="MIR repeat: matches 34..256 of consensus"
7203..7407
/note="MIR repeat: matches 2..208 of consensus"
complement(8312..8494)
/gene="dJ862P8.2"
complement(<8312..8494)
/gene="dJ862P8.2"
/product="dJ862P8.2 (5' end of gene encoding novel protein
similar to C.elegans and D.melanogaster)"
/evidence=not_experimental
complement(<8312..8464)
/gene="dJ862P8.2"
/note="match: proteins: Tr:046074"
/note="match: proteins: Tr:046074"
/codon_start=1
/evidence=not_experimental
/product="dJ862P8.2 (5' end of gene encoding novel protein
similar to C.elegans and D.melanogaster)"
/protein_id="CAC17570.1"
/db_xref="GI:11493212"
/translation="MVSQVQLRLRQGVMAALNGWYHDPQSKFTNSCHLYLMLFLL
LPALHL"
8840..9031
repeat_region
/note="96 copies 2 mer ct 70% conserved"
10240..10362
repeat_region
/note="L2 repeat: matches 2576..2702 of consensus"
10558..10791
repeat_region
/note="L2 repeat: matches 2157..2417 of consensus"
11593..11897
repeat_region
/note="Alusx repeat: matches 1..312 of consensus"
11942..12373

```

repeat_region      /note="L2 repeat: matches 1642. .2091 of consensus"
12715. .12791
/note="LIR33 repeat: matches 99. .175 of consensus"
repeat_region      13110. .13274
/note="LIM4 repeat: matches 4431. .4608 of consensus"
misc_feature        14285. .14806
/note="match: GSS: Em:AQ476745"
repeat_region      14369. .14623
/note="LIM4 repeat: matches 3588. .3847 of consensus"
repeat_region      14684. .14781
/note="HAL1 repeat: matches 842. .935 of consensus"
repeat_region      14782. .15078
/note="Alusx repeat: matches 1. .300 of consensus"
repeat_region      15079. .15181
/note="HAL1 repeat: matches 935. .1051 of consensus"
repeat_region      15294. .15791
/note="LIM4 repeat: matches 2865. .3375 of consensus"
repeat_region      16096. .16185
/note="LIM4 repeat: matches 2693. .2784 of consensus"
repeat_region      16241. .16506
/note="Alub repeat: matches 24. .288 of consensus"
repeat_region      16509. .16548
/note="20 copies 2 mer at 77% conserved"
repeat_region      16577. .16638
/note="LIM4 repeat: matches 2375. .2437 of consensus"
repeat_region      16639. .16726
/note="LIP4 repeat: matches 6057. .6144 of consensus"
repeat_region      16727. .16989
/note="LIM4 repeat: matches 2120. .2375 of consensus"
repeat_region      16742. .17063
/note="LIMEC repeat: matches 2407. .2364 of consensus"
repeat_region      18214. .18422
/note="LIMEC repeat: matches 2105. .2311 of consensus"
repeat_region      18423. .18740
/note="Alub repeat: matches 1. .309 of consensus"
repeat_region      18741. .19382
/note="LIMEC repeat: matches 2311. .2620 of consensus"
repeat_region      19580. .19641
/note="LIMB8 repeat: matches 6113. .6174 of consensus"
repeat_region      19776. .19948
/note="LIME3A repeat: matches 5980. .6160 of consensus"
repeat_region      20173. .20579
/note="MSTRB repeat: matches 1. .422 of consensus"
repeat_region      20631. .20842
/note="MIR repeat: matches 22. .245 of consensus"
repeat_region      21161. .21474
/note="Alusx repeat: matches 1. .312 of consensus"
repeat_region      21772. .21873
/note="51 copies 2 mer aa 60% conserved"
repeat_region      22246. .22531
/note="Alusx repeat: matches 1. .289 of consensus"
repeat_region      22582. .22894
/note="Alusx repeat: matches 1. .311 of consensus"
repeat_region      23101. .23421
/note="LIM4 repeat: matches 2740. .3046 of consensus"
repeat_region      23468. .23548
/note="MIR1J repeat: matches 104. .191 of consensus"
repeat_region      23643. .24023
/note="MSTRB repeat: matches 1. .418 of consensus"
repeat_region      24596. .24676

```

```

Query Match      83.2% Score 136.4; DB 9; Length 118632;
Best Local Similarity 99.3%; Pred. No. 1,1e-29;
Matches 137; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

Oy 27 TTTTGCTACTTGAGAGATAGAACATGATGACATCTGCAATTAACCTTGAAGATTACAG 86
|||||
Db 59428 TTTTGCTACTTGAGAGATAGAACATGATGACATCTGCAATTAACCTTGAAGATTACAG 59487
|||||
Oy 87 ATTTTGGGTTGGCGAGGAATGCGACAGACCAACCAAAATGAGCAGACAGGCACTTATG 146
|||||
Db 59488 ATTTTGGGTTGGCGAGGAATGCGACAGACCAACCAAAATGAGCAGACAGGCACTTATG 59547
|||||

```

```

Oy 147 CCTGATGCGCCCGAGAG 164
Db 59548 CCTGATGCGCCCGAGAG 59565

```

```

RESULT 8
BC021891
LOCUS
DEFINITION
Mus musculus, similar to mitogen-activated protein kinase
ACCESSION
BC021891
KEYWORDS
BC021891.1 GI:18257337
SOURCE
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

```

```

REFERENCE
1 (bases 1 to 3616)
AUTHORS
Strausberg,R.
TITLE
Submitted (18-JAN-2002) National Institutes of Health, Mammalian
JOURNAL
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

REMARK
NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT
Contact: MGC help desk
Email: cgaabs-remail.nih.gov
Tissue Procurement: Jeffrey Green M.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Huilyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>
Series: IRM Plate: 35 Row: d Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

```

FEATURES
Source
location/Qualifiers
1..3616
/organism="Mus musculus"
/db_xref="taxon:10090"
/map="FVB/N"
/clone="MGC:27778 IMAGE:3156324"
/tissue_type="Mammary tumor. C3(1) -Tag model. Infiltrating
ductal carcinoma. 5 month old virgin mouse."
/clone_lib="NCL CGAP Mam6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
224..3229
/codon_start=1
/product="Similar to mitogen-activated protein kinase
kinase kinase 9"
/protein_id="AAH21891.1"
/db_xref="GI:18257338"
/translat="MALPYAEGTADPLSPARDGSGTSSGMAALYDEARGEDELS
LRGQLEVEYLDSDAAVSGDEGMMAGOVORRLGIFPASVYAPGPPVPPPPRCSP
VHDFPERLEKELIGAGGGGOYRATWGOOEYAVKARRDPODAAAASVRRARL
FAMLRPNITQLRGVCLRDPHCIVLEFARGAALNLLAAASDPRAPRRARRIP
QVLVNNAVOIARGMILUBEAVYPIILHRDKSNLILLEKIHDDICNKTITPQGL
ARWMHRTTRASAGTYAAMAPEYTRSLTSSKSDTWSTGLVLMELTGEVPRGIDGL
AVAYGAVALNLTLPSTCPDEPFAKLKMKCWDDPIIRSFALIIOQLTALAEAVLTN
MPESEFHSMDKRLKLEIOMFSELTAKELKRSDEELSPAALOQSOEALLRBDQOL
AEREIVYLERELNVLIFOLSOEAPHYKRRKGRFRRLRLKDGHSILSPDOHRTV
QASPTLDKRSSSDSGICSPGSPIMLPRIRALIOIYSDENNKTRGNMVRRODFEDVK
RSFKKKGCTMGPSVOTKERPERREYRPLSDGNSWSSLIKSKTTPLAISLFDPQ

```

CDS

Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S.,
 Fero, S., Ferreira, P., FitzGerald, M., Gage, D., Galaan, J.,
 Gardyna, S., Gard, S., Graham, L., Grand-Pierre, N., Haeb, B.,
 Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Komel, A.,
 Karats, A., Kells, C., Landers, T., Levine, R., Lindblad, K.,
 Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C.,
 McCarthy, M., Meldrum, J., Meneus, L., Mihova, T., Mlenga, V.,
 Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H.,
 O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, J., Peterson, K.,
 Phunkhuan, P., Pierre, N., Raymond, C., Rella, R., Rise, C., Rogov, P.,
 Roman, J., Roy, A., Schauer, S., Schnuppach, R., Seaman, S., Severy, P.,
 Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas,
 Tefaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H.,
 Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zalnoun, J.,
 Zembek, L., Zimmer, A. and zody, M.
 Direct Submission
 Submitted (31-RNG-2002) Whitehead Institute/MIT Center for Genome
 Research 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 21, 2002 this sequence version replaced gi:17061506.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996 1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center Code: WTRR

```

Cook,A., Cooke,P., Dearellano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Garday,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrim,J., Menus,L., Milhova,T., Mlenga,V.,
Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norb,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
Plunkhamp,P., Pierre,N., Raymond,C., Reiter,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schuback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tasfaye,S., Theodores,J., Topham,K., Travers,M., Vasiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zalnoun,J.,
Zemke,L., Zimmer,A. and Zody,W.
Direct Submission
Submitted (21-Nov-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 21, 2002 this sequence version replaced gi:17061506.
All repeats were identified using RepeatMasker:
Smith, A.P.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
-----
Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence.submissions@genome.wi.mit.edu
-----
Project Information
Center project name: L18760
Center clone name: 553_O_21
-----
Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 161900 bases at least Q40
Consensus quality: 165140 bases at least Q30
Consensus quality: 166006 bases at least Q20
Insert size: 176000; agarose-fp
Insert size: 166395; sum-of-ctrls
Quality coverage: 5.7 in Q20 bases; agarose-fp
Quality coverage: 6.1 in Q20 bases; sum-of-ctrls
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 25 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1
* 894: contig of 894 bp in length
* 895 994: gap of 100 bp
* 995 2169: contig of 1175 bp in length
* 2170 2269: gap of 100 bp
* 2270 3068: contig of 799 bp in length
* 3069 3168: gap of 100 bp
* 3169 4507: contig of 1339 bp in length
* 4508 4607: gap of 100 bp
* 4608 5627: contig of 1020 bp in length
* 5628 5727: gap of 100 bp
* 5728 7115: contig of 1388 bp in length
* 7116 7215: gap of 100 bp
* 7216 8963: contig of 1748 bp in length
* 8964 9063: gap of 100 bp
* 9064 11712: contig of 2649 bp in length
* 11713 11812: gap of 100 bp
* 11813 14077: contig of 2265 bp in length
* 14078 14177: gap of 100 bp
* 14178 17221: contig of 3044 bp in length
* 17222 17321: gap of 100 bp
* 17322 19510: contig of 2189 bp in length
* 19511 19610: gap of 100 bp
* 19611 23807: contig of 4192 bp in length

```


Scherer, S., Scott, G., Shen, H., Shooshtrani, N., Sisson, I.,
Sodergren, E., Sonaike, T., Sparks, A., Stanley, H., Stone, H.,
Tatey, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinsón, R., Wang, Q.,
Wang, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczky, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.

Direct Submission
Unpublished
2 (bases 1 to 96101)
Worley, K.C.

Direct Submission
Submitted (19-FEB-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 96101)
Worley, K.C.

Direct Submission
Submitted (13-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 12, 2002 this sequence version replaced gi:18701953.

----- Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc.help@bcm.tmc.edu

----- Project Information
Center project name: G02X
Center clone name: CH230-234D24

----- Summary Statistics
Sequencing vector: Plasmid:
Chemistry: Dye-terminator Big Dye 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 49084 bases at least Q40
Consensus quality: 53702 bases at least Q30
Consensus quality: 57526 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 49 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1
1115 1114: contig of 1114 bp in length
1215 1214: gap of unknown length
1215 1214: gap of unknown length
2723 2722: contig of 1508 bp in length
2823 2822: gap of unknown length
4183 4182: contig of 1360 bp in length
4283 4282: gap of unknown length
5939 5938: contig of 1656 bp in length
6039 6038: gap of unknown length
7048 7047: contig of 1009 bp in length
7148 7147: gap of unknown length
8410 8409: contig of 1262 bp in length
8510 8509: gap of unknown length
9890 9889: contig of 1380 bp in length
9990 9989: gap of unknown length
11463 11462: contig of 1473 bp in length
11563 11562: gap of unknown length
12743 12742: contig of 1180 bp in length
12843 12842: gap of unknown length
13853 13851: contig of 1009 bp in length
13952 13951: gap of unknown length
15438 15437: contig of 1486 bp in length
15538 15537: gap of unknown length
17126 17126: contig of 1589 bp in length

17127 17226: gap of unknown length
17227 17226: contig of 1406 bp in length
18633 18732: gap of unknown length
18733 20346: contig of 1614 bp in length
20347 20446: gap of unknown length
20447 21605: contig of 1159 bp in length
21606 21705: gap of unknown length
21706 22949: contig of 1244 bp in length
22950 24436: contig of 1387 bp in length
23050 24436: gap of unknown length
24437 26037: contig of 1501 bp in length
24537 26137: gap of unknown length
26038 28251: contig of 2124 bp in length
26138 28361: gap of unknown length
28262 29946: contig of 1585 bp in length
28362 30046: gap of unknown length
29947 31352: contig of 1306 bp in length
30047 31452: gap of unknown length
31353 32528: contig of 1076 bp in length
31453 32528: gap of unknown length
32529 34558: contig of 1930 bp in length
32629 34658: gap of unknown length
34659 36487: contig of 1829 bp in length
34658 38910: contig of 2323 bp in length
36488 39010: gap of unknown length
38911 39010: gap of unknown length
39011 40486: contig of 1476 bp in length
40487 40586: gap of unknown length
40587 41854: contig of 1268 bp in length
41855 41954: gap of unknown length
41955 44349: contig of 2395 bp in length
44350 44449: gap of unknown length
44450 46817: contig of 2368 bp in length
46817 46917: gap of unknown length
46917 48441: contig of 1524 bp in length
48442 48541: gap of unknown length
48542 49639: contig of 1098 bp in length
49640 49739: gap of unknown length
49740 51158: contig of 1419 bp in length
51159 51258: gap of unknown length
51159 53003: contig of 1745 bp in length
51259 53103: gap of unknown length
53004 53104: gap of unknown length
53104 55784: contig of 2661 bp in length
55785 55884: gap of unknown length
55885 58448: contig of 2564 bp in length
58449 58548: gap of unknown length
58549 60250: contig of 1702 bp in length
60251 60350: gap of unknown length
60351 62642: contig of 2292 bp in length
62643 62742: gap of unknown length
62743 63903: contig of 1161 bp in length
63904 64003: gap of unknown length
64004 66598: contig of 2595 bp in length
66599 66698: gap of unknown length
66699 69090: contig of 2392 bp in length
69091 69190: gap of unknown length
69191 72504: contig of 3414 bp in length
72505 72704: gap of unknown length
72705 75514: contig of 2810 bp in length
75515 75614: gap of unknown length
75615 78089: contig of 2475 bp in length
78090 78189: gap of unknown length
78190 79571: contig of 1382 bp in length
79572 79671: gap of unknown length
79672 82016: contig of 2345 bp in length
82017 82117: gap of unknown length
82117 84149: contig of 2033 bp in length
84150 84249: gap of unknown length
84250 88193: contig of 3944 bp in length
88194 88293: gap of unknown length
88294 91731: contig of 3438 bp in length
91732 91831: gap of unknown length

FEATURES	Source	Location/Qualifiers
FEATURES	*	91832 96101: contig of 4270 bp in length.
Source		I . 96101
		/organism="Rattus norvegicus"
		/db_xref="taxon:10116"
		/clone="CH230-234D24"
BASE COUNT	23908 a 21808 c 22462 g 23062 t	4861 others
ORIGIN		
Query Match	63.7%; Score 104.4; DB 2; Length 96101;	
Best Local Similarity	84.8%; Pred. No. 3.2e-20;	
Matches 111; Conservative	0; Mismatches 21; Indels 0; Gaps 0;	
Oy	27	TTTTGCTACTGTGAGAAAGTAGACATGATGATGCATCTGCATTAACAACTTTGAGATTACAG 86
Db	33910	TTTTGCTCTCTGGAGAAAGTAGACACGATGACATTTTGCAACAAACCTTGAGATACAGG 33911
Oy	87	ATTTTGGTTGGCGAGGAGATGCGACAGACACCAAAATGAGCAGACAGACCTATG 146
Db	33910	ACTTCGGTGTGGCGAGGAGATGCGACAGACACCAAGATGAGCGAGCAGAACCTACG 33851
Oy	147	CCTGGATGGCCCCAGAG 164
Db	33850	CCTGGATGGCCTCGAG 33833
RESULT 11		
LOCUS	AF251442	
DEFINITION	Homo sapiens mixed lineage kinase MUK1 mRNA, partial cds.	
ACCESSION	AF251442	
VERSION	AF251442.1 GI:12005723	
KEYWORDS		
SOURCE	Homo sapiens.	
ORGANISM	Homo sapiens.	
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.	
REFERENCE	1 (bases 1 to 3931)	
AUTHORS	McNee,J.J., Dower,S.K. and Guesdon,F.	
TITLE	cDNA sequence and gene organisation of mixed lineage kinase 1	
JOURNAL	Unpublished	
REFERENCE	2 (bases 1 to 3931)	
AUTHORS	McNee,J.J. and Guesdon,F.	
TITLE	Direct Submission	
JOURNAL	Submitted (03-APR-2000) Molecular and Genetic Medicine, University	
	of Sheffield, Royal Hallamshire Hospital, Glossop Road, Sheffield	
	S10 2UF, United Kingdom	
FEATURES		
Source		1..3931
	/organism="Homo sapiens"	
	/db_xref="taxon:9606"	
	<1..3922	
	/note="protein kinase"	
	/codon_start=2	
	/product="mixed lineage kinase MUK1"	
	/protein_id="AA044591.1"	
	/db_xref="GI:12005724"	
	/translation="PLPYTWAVEFEYAGEDELTRLGADVENVLSKQSYSGDEGWT	
	GLNMQIIPSSNVYTPRPAFSRSCQGGEDPCYPPIQLLEIDFALTLEETIGIGG	
	FGKYRAFWIDEEVAARHDPDEDISQITENYRQEAFLFAMLKHNIITALRGVCLK	
	EPNCLVMEFARGGPLRNVRISGKRIIPDILVMAVVOIARGMNVLHDAIVIPRIHRLK	
	SNLITLLOKVENGLSNKILKTIDFEGILAREHMPITSSAAGTYKMADEVIRASMEFSK	
	GSIDWYGVILWELTLEGEVPRGIDGLAVLYGAMNMLALPIPTSCPEPAKIMEDCW	
	NPDHSRSPSTNITLDLUTTTIESGFPEFKPDSFICLDDNNKHLEQEMFDLRAKEKEL	
	RTWELTIRALDQKNODELLRRKEQLEAERDILERELENTIHQLCQEPVKKKK	
	GKPKSRILKLDGNRI SLPSDFORKEFTVQASPTMDKRSKLINSRSPASPTIIPRIAR	
	AIOITLPGESSKTSWGRSVYVKEEGEEDERAKPKKRWGSLGQKELASGEGSPQ	
	REFRANGSTPESPHFLIGLKSVDYGVKQWSSAPNIDKPRSSPALPGTTSLMEMAA	
	LLASWVYDIEEDDESEGGSGESRIHQSPSYLCIPPREGDDGSPSGIHREH	
	PTPNASATSTQILPTNYSIKRGAAHHRGCPVALIGCCAVLAATGLGFDILPAKCOLIL	
	PLDEEPAREEKKRRRGSRSPRSTSPSRKLPKKEEPMILAGDSASLITLIS	
	LSSTISECVNSTRSLRSDDELIVYEMPVSEAPLSPCHNPLVNVVEREKKDPVQ	
	SLTPHYVITTPSQPSHRRPDSGALKPTELIALSRSPSSNGSLSPGAGALTKPSPS	

[illegible]

BASE COUNT	632 a	1207 c	1143 g	549 t
ORIGIN	AVOAAGMHYHCALCVPIHRLDKSNNTLLIQIESNDMHKTLKLTIDGIAAEHMK TTKSAGATYAMMADEVYKASTSEKSGDWSFGVLLDELLGEPTPYGIDCLAYATV AVNKLTLPISPTCEPPPAQIADWCADDPHRPDPFASIIQQLEALAEQVIREMPDSEF HSMQEGMKRLQIGLFDLRAKEXELLSREBELTEAAEORNSOALRREHLLAOMEL EYFERRELTLLQIQDREDPHRYRRRGTFEASKSLARGGEEISMLDLFKRHTVQASB GLDDRNVYEVYEGPDSPLTPPEFRALIOLEPAVGGAMROSPPRLSDSNERRACWAM GSSPKRPGEAQNGRRSRMDPATWYLLDSDSPSGSPTSPPALNGNPPRPSLEPEEK RPYVAERKSSSGTPKLQIALRLRGTALLASLGSRDLQPPGPERGESSPTPTPTPT PAPCPTEPSPPLICFLSKTPTDSPPTPAPLLDLIDGIVGONASASPRREEPRGCTVTS PPTPTSSAPTEPCTPRSPPLGLISRPSPSLRIRIDPMSVSGSPSPSPSPQAPAP RRAWTLFPDDSDPMDSPAPNPFQGGQDCCRAQKDMGAQAPWVEAGP"			
misc_feature	577.. 756			
	/gene="sprk"			
	/note="SH3 domain"			
misc_feature	793.. 1659			
	/gene="sprk"			
	/note="kinase domain"			
misc_feature	1710.. 1896			
	/gene="sprk"			
	/note="leucine/Isoleucine zipper region"			
polyA_site	3531			
	/gene="sprk"			
	/note="32 A residues"			

Query Match	50.6%	Score 83	DB 9	Length 3531
Best Local Similarity	69.3%	Pred. NO. 5.6e-14		
Matches 113	Conservative 0	Mismatches 50	Indels 0	Gaps 0

QY	2	CACGGGACATCAAGGCGGAAATTTTGGTACTGTGGAAGATAGAAACATGATGATC	61
Db	1165	CACGGTATCTCAAGTCCACACACATTTTGGTGTGCGACCCATTGTAGAGTGCACATG	1224
QY	62	TGCATAAACTTTGGAATATACAGATTTTGGGTTGGCGAGGGAATGCGACAGACACC	121
Db	1225	GAGCACAAGACCTCTBMAATCACCGACTTTTGGCTTGGCCGCGAGAGTGGCACAAAAACCA	1284
QY	122	AAATGAGCACAGCAGGCACTTATGCTGGATGGCCCCAGAAG	164
Db	1285	CAATATGATGCGCGGGGACCTACGGCTCGATGTGGCTCTCTTGAGG	1327

RESULT	13			
AX399680				
LOCUS	AX399680	3558 bp	DNA	linear
DEFINITION	Sequence 1 from Patent WO022947.			
ACCESSION	AX399680			
VERSION	AX399680.1	GI:21335453		
				PAT 06-JUN-2002

SOURCE	human.
ORGANISM	Homo sapiens Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE	1
AUTHORS	Delaney,A.D. and Yoganathan,T.
TITLE	Cancer associated protein kinases and their uses
JOURNAL	Patent: WO 0224947-A 1 28-MAR-2002.
FEATURES	KINETEK PHARMACEUTICALS INC (CA); UNIV BRITISH COLUMBIA (CA) Location/Qualifiers

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
483.3025
CDS

```

```
/translation="MEPIKSLFLKPSLSPGWMNWSGSGGPGGPGGPGKAAAGTAV
VWMLADPVEPSGQDDELAKRQDRVEVLSIDMISGDEGMMAGGVOGVOGTPSNYSR
GGGPBPCEVAFROELRLEFVYIGCGRGKRYKRSGRKRELYAVKARADPPEDISVTAES
VREAEALFAMLAHPNITIALKACGLEEPNCLVETAVAGSPGLSTALGRPVPVAVLWV
AVOIAAGMHLCEALPVIHRDGKSNNTLLLOPTESDMEKHTLKTIDFGCEARWK
```

TTQMSAGTYANMAPEVIYKASTFESKSDVMSFVYLLMELLTGEVPHYGIDCLAVATCV
AVNNLITLIPSTCEPRPEPFASIILOOLEMADWADPPDPRPFASIILOOLEMAYLRNEMRDSF
HSMGEWKEIREIOGLFEDBLRAKECMLLSREBELTAAAEORSQAOLNRRHLLAOMEL
EYFERRELTLLLOQVDRRPHVRRRRGTTFKPKLAARGOEKISMPIDPKRITVOASP
GLDRRRNVEYFGPDSPTPFPRFRAIOLTEPAPGAMGORSRPSITLDSNBRRCWAM
GPSSPGEGEAONGRRSRMDATWYLLDSBSPDGLSGSTPPLNAGNPPRSLIEEERK
RVAERGGSSSGTKLJORALLRGTTALLASIGLGRDLOPGGPGREGGSGSPTRPTPT
PAPCPTEPPSPCLTCSLKTDPDSPPTPRAITLDIGIYVQORSKASPRREDEPRGTYTS
PPRPTSASAPETPCTPPSPPLGLISRRPSPSLRNDPMSVSGAPSPSLPSPPOAP
RRAWITLPPDSDPMDSPPANPFGGGGDDCAQTKRQDGAQAPVWPEAG"

BASE COUNT
ORIGIN

635 a 1211 c 1161 g 551 t

Query Match	50.68;	Score	83;	DB	6;	Length	3558;
Best Local Similarity	69.38;	Pred.	No.	5.0e-14;			
Matches	113;	Conservative	0;	Mismatches	50;	Indels	0;
						Gaps	0;

0Y 2 CACCGGACATCAAGCAGGAATATTTGCTACTTGAAGATAGAACATGATGCATC 61
 Db 1196 CACCGTATCTAAGTCCAAACACATTTTGCTGCTGCAGCCATTGAGAGTACGACGAT 1255

0y 62 TGCATTAACCTTTGAAGATTACAGATTTTGGGTTGGCGAGGAATGACACAGGACC 121
Db 1256 GACGACACGACCCTGAGAGATCACCGACTTTGGCTGGCCCGGAGAGTGCACAAACCCACA 1315

QY	122	AAATGAGCACACGACGACCTATGCTGGATGGCCCCAAG	164
Db	1316	CAATGATGTCGCGGGCACTACGCTGAGAGGCTCTGAGG	1358

RESULT 14	
LOCUS	HUMMLK3A
DEFINITION	Human protein kinase (MLK-3) mRNA, complete cds.
ACCESSION	L32976
VERSION	L32976.1 GI:488295
KEYWORDS	MLK-3 gene; protein kinase.
SOURCE	Homo sapiens (tissue library: lambda g10) thymus cDNA to mRNA.
ORGANISM	Homo sapiens
REFERENCE	Emkaroyita; Metazoa; Chordata; Craniata; Vertebrata; Eucleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS	1 (bases 1 to 3558)
TITLE	Ing, Y. L., Leung, I. W., Heng, H. H., Tsui, L. C. and Lassem, N. J.
JOURNAL	MLK-3: Identification of a widely expressed protein bearing
MEDLINE	an SH2 domain and a leucine zipper-basic region domain
PUBMED	Oncogene 9 (6), 1745-1750 (1994)
	94239754
	8183572

```

gene
  /organism="Homo sapiens"
  /db_xref="taxon:9606"
  /map="11 q13.1-13.3"
  /tissue_type="thymus"
  /tissue_1fb="lambda gt10"
  1..3558
  /gene="MLK-3"
  482..3025
CDS

```

```

/standard_name="PTK-1, sprk"
/note="amino acid feature: SH3 domain, bp 613 . . 789;
amino acid feature: leucine zipper basic region, bp 1686
. . 1932; amino acid feature: proline-rich region, bp 2466
2839"

```

[illegible]

TTOMSACTIANNAAEVIATSAFFSKSDSWSEGVLLMELLETIGVPRQIDCLAVIAYE
AVNKTLTIPSCPPPEPQALMDCAMDQHDWRDPDSLLOOLEALAEQVLRWIDSEFSE
ISMQDGRKEITDGLDELBAKREKELLREEBUTRARBEORSQAOLARRHLLAQMEL
GLDERNNVEVWPGSPPTPRRALOLEAPBOADGSGRIEMPDLEKRIITVOVASQ
GPSSEPQEAQGRSNDMETATLSDSPSTSGSPSTPLALNGPVRSLIEEBEKEP
RVPAPERSSESTPKLLOALLRGWLLASLISGLDLOPPGQGRGEGSEPTPLPPPT
PAPCETPEPPSPLLTFSJLTQDPSPPRAUJLLDGLTYQGRSANKSPREPEEPQYSL
PQVPSKNAPEVGTIPKPSLGLLISNPSPLRSJLDIPWYSKAPVAPNGP
BRAPPTLPDSDPPEPAPNFOGSGPQOCRAQTCKMGAAAPVPEAGP"

